Reserverechnung nach Lidstone unter Verwendung eines durch Stichproben bestimmten durchschnittlichen Schlussalters

Inaugural-Dissertation

zur Erlangung der philosophischen Doktorwürde
vorgelegt der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Basel

von

Robert Baumann
aus Basel

Bern · Gedruckt bei Stämpfli & Cie · 1950

Reserverechnung nach Lidstone unter Verwendung eines durch Stichproben bestimmten durchschnittlichen Schlussalters

Inaugural-Dissertation

zur Erlangung der philosophischen Doktorwürde
vorgelegt der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Basel

von

Robert Baumann
aus Basel

Bern · Gedruckt bei Stämpfli & Cie · 1950

Genehmigt von der Philosophisch-naturwissenschaftlichen Fakultät auf Antrag der Herren Prof. Dr. E. Zwinggi und P. D. Dr. M. Gürtler.

Basel, den 4. Juli 1950.

R. Wizinger, Dekan

R. UNIV.
BIBLIOTHEEK
LEIDEN

Meinem Vater gewidmet

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41544315_0072

Inhaltsverzeichnis

Seite

I. Abschnitt: Lidstoneverfahren, Problemstellung 7

II. Abschnitt: Theorie der Stichproben 9

§ 1. Verteilungen und einfache Prüfverfahren

a) Die Normalverteilung 9

b) Die χ^2-Verteilung 10

c) Die t-Verteilung 11

d) Die f-Verteilung 12

§ 2. Weitere Prüfverfahren

a) Prüfung auf Normalverteilung mit γ 13

b) Prüfung auf Normalverteilung mit χ^2 14

c) Prüfung auf Homogenität 15

III. Abschnitt: Untersuchung von Stichproben auf Normalverteilung . . . 18

IV. Abschnitt 28

§ 1. Durchschnittliche Schlussalter und Reserven eines effektiven Versicherungsbestandes 28

§ 2. Untersuchung auf Homogenität 30

§ 3. Konstruktion eines theoretischen Bestandes 33

§ 4. Wahrscheinlichkeiten für Reserveabweichungen 35

§ 5. Versuch aus der Volksversicherung 41

V. Abschnitt: Versuch aus der Grosslebensversicherung 43

Zusammenfassung 46

Literaturverzeichnis 47

Lebenslauf 48

I. Abschnitt

Lidstoneverfahren [1], Problemstellung

Ein Bestand von gemischten Versicherungen werde derart aufgeteilt, dass in eine Gruppe jeweilen lauter Versicherungen mit gleichem Ablaufjahr fallen und demnach die gleiche künftige Dauer $m = x + n - z$ aufweisen. (Es bedeutet dabei $z = x + t$ das erreichte Alter eines Versicherten.) Nach dem Näherungsverfahren von *Lidstone* ist alsdann das Nettodeckungskapital der Gruppe gegeben durch

$$V \approx A_{s-m:\overline{m}|} \sum_i S_i - a_{s-m:\overline{m}|} \sum_i S_i P_{x_i},$$

wobei s den Sinn eines durchschnittlichen Schlussalters der Gruppe hat und aus

$$c^{s-50} = \frac{\sum_i Z_i}{\sum_i S_i}$$

zu bestimmen ist. S_i bedeutet die Versicherungssumme und P_{x_i} die Nettoprämie für die Leistung «1». Die Grösse

$$Z_i = S_i \, c^{z_i + m - 50},$$

der sog. *Lidstonekoeffizient*, ist eine von der Dauer t unabhängige Hilfszahl und kann für jede einzelne Versicherung ein für allemal bestimmt werden. (Dabei ist c^{-50} lediglich eingeführt worden, um keine allzu grossen Zahlenwerte zu erhalten.)

Wir erkennen aus der Reservegleichung, dass es nötig ist, neben der Versicherungssumme und der Nettoprämie auch den Lidstonekoeffizienten für die einzelne Versicherung zu haben und in jeder

[1] Die Ziffern in eckigen Klammern beziehen sich auf die Literaturangaben am Schlusse der Arbeit.

Gruppe (= Ablaufjahr) zu summieren. Diese Feststellung führt fast notwendig auf die Frage, ob es nicht möglich wäre, das Verfahren derart zu vereinfachen, dass auf das Fortschreiben der Lidstonekoeffizienten verzichtet und auf ein aus Stichproben bestimmtes Schlussalter abgestellt wird. Diese Möglichkeit ist bisher nicht geprüft worden, wie uns überhaupt auf dem ganzen Gebiete der Reserverechnung aus Stichproben nur drei Arbeiten bekannt sind [2]. Das Problem ist aber ganz bestimmt wert, genauer untersucht zu werden; dies um so mehr, als ohne mathematische Ableitung erkennbar ist, dass das Verfahren von *Lidstone* sich besonders gut für die Bestimmung der Hilfszahl aus Stichproben eignet. Der Grund ist der folgende: In den Ablaufjahren mit kurzer künftiger Dauer treffen die im Verhältnis zur Summe grössten Reserven in der Regel mit den grössten Versicherungssummen zusammen. In den gleichen Jahren wirkt sich aber ein Fehler in der stichprobenweisen Bestimmung des durchschnittlichen Schlussalters nicht stark aus, weil die Reserve ohnehin nicht mehr stark vom Alter abhängt. Erst mit wachsendem m erhält eine Abweichung immer mehr Gewicht, so dass auch der relative Unterschied zwischen genauer Reserve und Stichprobenreserve dauernd zunimmt. Gleichzeitig sind aber diese Reserven im Verhältnis zur Reserve des Gesamtbestandes sehr klein, und damit sind es auch die absoluten Abweichungen. So können wir zusammenfassend sagen, dass auch bei einer ungenauen Bestimmung des durchschnittlichen Schlussalters der Fehler in der Gesamtreserve nicht allzu gross werden kann.

Mit dieser Erkenntnis können wir die Frage wie folgt aussprechen: Wie verhält sich der relative Reservefehler, der mit einer bestimmten Wahrscheinlichkeit nicht überschritten wird, wenn zur Bestimmung des Schlussalters dem Bestande eines jeden Ablaufjahres eine vorgeschriebene Anzahl Stichproben entnommen wird?

Bei der Behandlung der gestellten Frage haben wir uns bewusst zu sein, dass eine allgemeingültige Antwort kaum jemals wird gegeben werden können, denn wir werden uns für die Bildung von Kriterien immer auf empirische Werte aus Versuchsbeständen stützen müssen. Das Ergebnis trifft für den gerade untersuchten Bestand zu, darf aber nur sehr vorsichtig auf andere Bestände übertragen werden. Auch haben wir zu beachten, dass bei den Versuchen wohl kaum jemals alle Bedingungen erfüllt sein werden, die von der strengen Theorie verlangt werden.

II. Abschnitt

Theorie der Stichproben [3]

§ 1. Verteilungen und einfache Prüfverfahren

a) Die Normalverteilung [3, 4]

Bei Vorliegen von Normalverteilung ist die Wahrscheinlichkeit, dass der Wert zwischen x und $x + dx$ liegt, gleich

$$dF = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx .$$

Dabei bedeutet μ das Mittel und σ^2 die Streuung um das Mittel. Es gibt sodann $\int_a^b dF$ die Häufigkeit der Beobachtungswerte im Intervall a, b an. Wird der Gesamtinhalt der Fläche unter der Häufigkeitskurve als Einheit gewählt, dann stellt ein bestimmter Teil dieser Fläche die relative Häufigkeit dar und gibt die Wahrscheinlichkeit dafür, dass eine Beobachtung in das Intervall a, b fällt. Es wird aber dann $\int_{-\infty}^{\infty} dF = 1$, wenn wir die Abweichung vom Mittel mit der Streuung als Einheit messen (standardisierte Normalverteilung), d. h. $\xi = \frac{x - \mu}{\sigma}$ setzen. Für den späteren Gebrauch können wir damit folgendes festhalten: Wird für eine Verteilung von Stichproben die zugehörige theoretische Normalverteilung berechnet, so brauchen wir bloss zur standardisierten Normalverteilung überzugehen, um nach Multiplikation der dann für ein bestimmtes Intervall gegebenen relativen Häufigkeit [5] mit N, d. h. mit dem Umfang der Verteilung, die Anzahl Beobachtungen im betreffenden Intervall zu erhalten. Umgekehrt können wir denjenigen Wert von ξ suchen, für den $\int_{-\infty}^{\xi} dF$ einer vorgegebenen relativen Häufigkeit H entspricht. Wählen wir etwa $H = 0{,}01$, dann wird $\xi = -2{,}326$; das bedeutet, dass nur 1 % aller Beobachtungen im Intervall $-\infty$ bis $\xi = -2{,}326$ liegen. Berücksichtigen wir weiter die Symmetrie der Normalverteilung, ferner den Umstand, dass für das Intervall $-\infty$ bis 0 gilt $H = 0{,}5$, so liegen wiederum 1 % aller

Beobachtungen im Intervall $\xi = 2{,}326$ bis ∞. Wir können damit zusammenfassend sagen, dass nur mit einer Wahrscheinlichkeit von 2 % Beobachtungen ausserhalb der Grenzen $\xi = \pm 2{,}326$ aus einer Normalverteilung entstanden sein können, was in der Regel nicht mehr als zufällige Abweichung gilt.

Ferner gilt folgende Feststellung: Wird einer grossen Grundgesamtheit eine genügend grosse Anzahl ausreichend grosser Stichproben entnommen und jeweilen das Mittel berechnet, so sind diese Mittel angenähert normal verteilt. Wenden wir diese Erkenntnis auf unsere Aufgabe an, so wird sie besagen, dass die $\bar{c} = c^{s-50} = \dfrac{\sum_i Z_i}{\sum_i S_i}$ angenähert normal verteilt sein werden, sofern wir einem Versicherungsbestande Stichproben von genügend grossem Umfang entnehmen. Es bleibt dabei späterer Abklärung vorbehalten, den Umfang der Stichproben festzulegen, so dass die $\bar{c}$ als normal verteilt angesehen werden können.

b) Die χ^2-Verteilung [3, 6]

Wir betrachten die Verteilung von

$$\chi^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{\sigma^2},$$

wobei wir $\mu = 0$ und $\sigma^2 = 1$ setzen und die x_i als normal verteilt voraussetzen. So vereinfacht sich die Aufgabe auf die Prüfung der Verteilung von $\chi^2 = \sum_{i=1}^{n} x_i^2$. Die Wahrscheinlichkeit dF, dass $\sum_{i=1}^{n} x_i^2$ zwischen χ^2 und $\chi^2 + d\chi^2$ liegt, ergibt sich zu

$$dF = \frac{1}{\left(\frac{n-2}{2}\right)!} \left(\frac{\chi^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{\chi^2}{2}} d\left(\frac{\chi^2}{2}\right).$$

Lassen wir n gegen ∞ gehen, so geht die χ^2-Verteilung in eine Normalverteilung über mit dem Mittelwert $\mu = \sqrt{n-1}$ und der Streuung $\sigma^2 = \frac{1}{2}$. Die Grösse $\sqrt{2\chi^2}$ ist alsdann um das Mittel $\mu = \sqrt{2n-2}$ normal verteilt mit der Streuung $\sigma^2 = 1$, was vor allem für die kaum mehr tabellierten Wahrscheinlichkeiten für $n > 30$ sehr gute und praktische Näherungswerte gibt.

c) Die t-Verteilung [3, 7]

Zur Prüfung der Abweichung des Mittelwertes $\bar{x}$ aus n' Werten vom Mittel μ dient der Ausdruck, sofern es sich um grosse Stichproben handelt und unter der Annahme, die Streuung σ^2 sei bekannt,

$$x = \frac{(\bar{x} - \mu)}{\sigma} \sqrt{n'}.$$

Dabei bedeutet x die Abszisse der standardisierten Normalverteilung, für welche wir die Wahrscheinlichkeit eines noch zufälligen Überschreitens angeben können. Sofern die Streuung σ^2 nicht bekannt ist, ersetzen wir sie durch die Streuung s^2 der Stichproben,

$$s^2 = \frac{\sum_{i=1}^{n'} (x_i - \bar{x})^2}{n' - 1}$$

und bestimmen als neue Masszahl

$$t = \frac{(\bar{x} - \mu)}{s} \sqrt{n' - 1}.$$

Dabei haben wir kein Recht mehr, anzunehmen, dass t noch normal verteilt sein wird und wir damit die Abweichung des Mittelwertes mit den Wahrscheinlichkeiten der Normalverteilung prüfen können. Die Verteilung von t hat folgende Gestalt

$$dF = \frac{\left(\frac{n-1}{2}\right)!}{\left(\frac{n-2}{2}\right)! \sqrt{\pi n}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt \quad \text{mit } n = n' - 1.$$

Die Verteilung von t ist symmetrisch, verläuft indessen flacher als die Normalverteilung und gibt darum für gleiche Abszissen grössere Wahrscheinlichkeiten eines noch zufälligen Abweichens. Bei grossem n lässt sich die t-Verteilung in die Normalverteilung überführen.

d) Die f-Verteilung [3c]

Wir nehmen an, wir hätten einer normalen Grundgesamtheit zwei Stichproben von je n_1' und n_2' Elementen entnommen. Als Streuung der beiden Stichproben hat man

$$s_1^2 = \frac{\sum_{i=1}^{n_1'} (x_i - \bar{x})^2}{n_1' - 1} \quad \text{und} \quad s_2^2 = \frac{\sum_{i=1}^{n_2'} (x_i - \bar{x})^2}{n_2' - 1}.$$

Die Frage geht nun dahin, ob man sich die beiden Streuungen s_1^2 und s_2^2 als durch Zufall aus der Streuung σ^2 einer einzigen Grundgesamtheit hervorgegangen denken kann, oder aber, ob zwei verschiedene Streuungen σ_1^2 und σ_2^2 zugrunde liegen. Man könnte zeigen, dass

$$(n_1' - 1)\frac{s_1^2}{\sigma^2} = \chi_1^2 \quad \text{und} \quad (n_2' - 1)\frac{s_2^2}{\sigma^2} = \chi_2^2$$

je einer χ^2-Verteilung gehorchen. Wir setzen $n_1' - 1 = n_1$ und $n_2' - 1 = n_2$, wobei wir n_1 und n_2 die Freiheitsgrade der Verteilung nennen wollen, und bilden

$$f = \frac{\chi_1^2 n_2}{n_1 \chi_2^2} = \frac{s_1^2}{s_2^2}.$$

Für die Wahrscheinlichkeit zufälliger Abweichungen der beiden Streuungen suchen wir demnach die Verteilung von f. Wir finden die gesuchte Verteilung, indem wir das Produkt bilden aus

$$dF_1(\chi_1^2) = \frac{1}{\left(\frac{n_1 - 2}{2}\right)!} e^{-\frac{\chi_1^2}{2}} \left(\frac{\chi_1^2}{2}\right)^{\frac{n_1-2}{2}} d\left(\frac{\chi_1^2}{2}\right)$$

und

$$dF_2(\chi_2^2) = \frac{1}{\left(\frac{n_2 - 2}{2}\right)!} e^{-\frac{\chi_2^2}{2}} \left(\frac{\chi_2^2}{2}\right)^{\frac{n_2-2}{2}} d\left(\frac{\chi_2^2}{2}\right)$$

und dann über χ_1^2 und χ_2^2 integrieren,

$$dF = \frac{1}{\left(\frac{n_1-2}{2}\right)!\left(\frac{n_2-2}{2}\right)!}\int\int\left(\frac{\chi_1^2}{2}\right)^{\frac{n_1-2}{2}}\left(\frac{\chi_2^2}{2}\right)^{\frac{n_2-2}{2}}e^{-\frac{\chi_1^2+\chi_2^2}{2}}d\left(\frac{\chi_1^2}{2}\right)d\left(\frac{\chi_2^2}{2}\right)$$

$$= \frac{\left(\frac{n_1+n_2-2}{2}\right)!\, f^{\frac{n_1-2}{2}}\, n_1^{\frac{n_1}{2}}\, n_2^{\frac{n_2}{2}}}{\left(\frac{n_1-2}{2}\right)!\left(\frac{n_2-2}{2}\right)!\,(n_2+n_1 f)^{\frac{n_1+n_2}{2}}}\,df.$$

§ 2. Weitere Prüfverfahren

a) Prüfung auf Normalverteilung mit γ [8]

Wir untersuchen eine Verteilung von Mittelwerten, angeordnet in Klassen x, auf Normalverteilung hin. Ist $f(x)$ die Besetzungszahl der Mittelwerte in einer Klasse, so gilt für die Gesamtzahl n der vorhandenen Werte $n = \sum_{i=1}^{n} f(x_i)$. Für eine dermassen angeordnete Verteilung bestimmen wir die folgenden Masszahlen:

$$k_1 = \frac{1}{n}\sum_{i=1}^{n} f(x_i)\, x_i, \qquad k_2 = \frac{1}{n-1}\sum_{i=1}^{n} f(x_i)\,(x_i - k_1)^2,$$

$$k_3 = \frac{n}{(n-1)\,(n-2)}\sum_{i=1}^{n} f(x_i)\,(x_i - k_1)^3.$$

Schreiben wir für

$$\sum_{i=1}^{n} x_i^k f(x_i) = S_k,$$

so können wir die k_i umformen zu

$$k_1 = m_1, \qquad k_2 = \frac{n}{n-1}m_2, \qquad k_3 = \frac{n^2}{(n-1)\,(n-2)}m_3,$$

wobei dann gilt,

$$m_1 = \frac{s_1}{n}, \qquad m_2 = \frac{s_2}{n} - m_1^2, \qquad m_3 = \frac{s_3}{n} - 3m_1 m_2 - m_1^3.$$

Die m_i entsprechen hier den Momenten der Normalverteilung. Wir betrachten sodann die Grösse

$$\gamma = \sqrt{\frac{(n-1)(n-2)}{6n}} \frac{k_3}{k_2 \sqrt{k_2}}.$$

Die Verteilung von γ ist symmetrisch, hat also den Mittelwert null. Das zweite Moment ist [9]

$$m_2(\gamma) = 1 - \frac{6}{n} + \frac{22}{n^2} - + \ldots$$

und das vierte Moment [10]

$$m_4(\gamma) = 3 - \frac{1056}{n^2} + \frac{24\,132}{n^3} - + \ldots,$$

während $m_3(\gamma)$, wie $m_1(\gamma)$ null ist. Mit wachsendem n geht das zweite Moment gegen 1, das vierte Moment gegen 3, während das erste und das dritte Moment verschwinden. Das sind aber die wohlbekannten Kriterien einer Normalverteilung, so dass wir für unsere Fälle die Masszahl γ als normal verteilt mit dem Mittelwert 0 und der Streuung 1 ansehen können. Wir haben damit aber auch ein brauchbares Kriterium für die Verteilung von Mittelwerten aus Stichproben, denn wir können für γ einen Schwellenwert ansetzen, der nicht überschritten werden darf, um noch als zufällig aus null entstanden angesehen zu werden. Setzen wir für unsere praktischen Versuche $\gamma = \pm 1{,}96$, so bedeutet das, dass nur in 5 von 100 Fällen γ eine grössere Abweichung erreichen könnte. Überschreitet γ diesen Wert, so zweifeln wir berechtigterweise, ob die Verteilung normal ist. Liegt γ aber innerhalb der Schwellenwerte, so können wir Normalverteilung annehmen, müssen uns aber bewusst sein, dass das Prüfverfahren mit γ vor allem auf die Symmetrie einer Verteilung anspricht.

b) Prüfung auf Normalverteilung mit χ^2 [11]

In unserem speziellen Falle bedeutet das Kriterium χ^2 eine Methode, zu entscheiden, mit welcher Wahrscheinlichkeit eine beobachtete Verteilung von Mittelwerten aus Stichproben noch durch Zufall aus einer

durch μ und σ^2 bestimmten theoretischen Normalverteilung hervorgegangen sein könnte. Wir nehmen an, $a_1, a_2, a_3, \ldots a_n$ seien die beobachteten, $\alpha_1, \alpha_2, \alpha_3, \ldots \alpha_n$ die entsprechenden theoretischen Häufigkeiten. Bilden wir die Masszahl

$$\xi^2 = \sum_{i=1}^{n} \frac{(a_i - \alpha_i)^2}{\alpha_i},$$

dann ist ξ^2 verteilt wie χ^2 mit $(n-1)$ Freiheitsgraden, wenn n die Anzahl Klassen bedeutet, wobei $P = \int_{\chi^2}^{\infty} dF(\chi^2)$ aussagt, mit welcher Wahrscheinlichkeit eine grössere Abweichung zwischen beobachteter und theoretischer Verteilung zu erwarten ist. Wir wollen auch bei diesem Verfahren als Grenze $P = 0{,}05$ annehmen, d. h., dass nur in 5 von 100 Fällen grössere Abweichungen noch durch Zufall aus einer Normalverteilung entstanden sein könnten. Für $P < 0{,}05$ zweifeln wir berechtigterweise, ob die zu prüfende Verteilung auch wirklich normal verteilt ist.

Wir haben angenommen, dass die theoretische Verteilung durch μ und σ^2 gegeben ist. Im konkreten Falle trifft das aber nicht zu, denn diese beiden Grössen müssen wir aus der Beobachtungsreihe als $\bar{x}$ und s^2 zuerst berechnen. Deshalb verringert sich die Zahl der Freiheitsgrade um 2, so dass ξ^2 verteilt ist wie χ^2 mit $(n-3)$ Freiheitsgraden.

c) Prüfung auf Homogenität [12]

Aus i verschiedenen Grundgesamtheiten entnehmen wir je eine Stichprobenreihe zu k Elementen, die wir folgendermassen gruppieren wollen:

$$\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & \ldots & x_{1k} \\ x_{21} & x_{22} & x_{23} & \ldots & x_{2k} \\ \vdots & & & & \\ x_{i1} & x_{i2} & x_{i3} & \ldots & x_{ik} \end{array}$$

Es sei $\bar{x}_i$ das Mittel der i-ten Stichprobenreihe und $\bar{x}$ das Mittel aus sämtlichen Stichproben. Die Gesamtzahl der Stichproben sei N. Weichen die $\bar{x}_i$ nicht wesentlich voneinander ab, dann können wir mit guten Gründen annehmen, dass die einzelnen Stichprobenreihen

aus Grundgesamtheiten entnommen worden sind, die miteinander identisch sind, wenigstens was den Mittelwert betrifft.

Zur Prüfung verwenden wir ein Verfahren, das sich auf einen in der Theorie der mathematischen Statistik bekannten Satz stützt, nämlich, dass sich eine Summe von quadratischen Abweichungen um ein Mittel einer Beobachtungsreihe zerlegen lässt in die Summe zweier unabhängiger Quadratsummen, von denen jede, wenn sie durch ihre entsprechende Anzahl Freiheitsgrade dividiert wird, eine Schätzung der Streuung der Grundgesamtheit darstellt.

Für unsere Darstellung der Mittelwerte gilt folgende Identität:

$$\sum_{ik} (x_{ik} - \bar{x})^2 = \sum_{ik} (x_{ik} - \bar{x}_i)^2 + \sum_{ik} (\bar{x}_i - \bar{x})^2,$$

oder, wenn n_i die Anzahl der Glieder in der i-ten Reihe bedeutet,

$$\sum_{ik} (x_{ik} - \bar{x})^2 = \sum_{ik} (x_{ik} - \bar{x}_i)^2 + \sum_{i} n_i (\bar{x}_i - \bar{x})^2.$$

Vorerst wollen wir annehmen, dass jede Grundgesamtheit normal verteilt sei mit der gleichen Streuung ν^2. Dann wollen wir voraussetzen, die Grundgesamtheiten hätten auch noch ein gemeinsames Mittel m.

Es gelten dann folgende Beziehungen:

Die Summe $\quad \frac{1}{\nu^2} \sum_{ik} (x_{ik} - \bar{x})^2$

ist verteilt wie χ^2 mit $\sum_i n_i - 1 = N - 1$ Freiheitsgraden.

Die Summe $\quad \frac{1}{\nu^2} \sum_{ik} (x_{ik} - \bar{x}_i)^2$

ist verteilt wie χ^2 mit $N - i$ Freiheitsgraden.

Die Summe $\quad \frac{1}{\nu^2} \sum_{i} n_i (\bar{x}_i - \bar{x})^2$

ist verteilt wie χ^2 mit $i - 1$ Freiheitsgraden.

Wir können diese drei Summen in folgendes Schema bringen, wobei wir auf die weitere Mitnahme von ν^2 verzichten, da es sich später herausheben wird:

Summen der Quadrate		Freiheitsgrade	Quotient	
zwischen dem Mittel einer Stichprobenreihe und dem Gesamtmittel	$\sum_i n_i (\bar{x}_i - \bar{\bar{x}})^2$	$i-1$	$\frac{1}{i-1} \sum_i n_i (\bar{x}_i - \bar{\bar{x}})^2$	(1)
zwischen den Stichproben einer Reihe und dem Mittel der Stichprobenreihe	$\sum_{ik} (x_{ik} - \bar{x}_i)^2$	$N-i$	$\frac{1}{N-i} \sum_{ik} (x_{ik} - \bar{x}_i)^2$	(2)
zwischen allen Stichproben und dem Gesamtmittel	$\sum_{ik} (x_{ik} - \bar{\bar{x}})^2$	$N-1$	$\frac{1}{N-1} \sum_{ik} (x_{ik} - \bar{\bar{x}})^2$	(3)

Wir halten dabei fest, dass die ersten beiden Kolonnen additiv verbunden sind, dass es also genügt, (1) und (3) zu berechnen, um (2) zu erhalten. In den späteren Versuchen nennen wir der Einfachheit halber (1) = «Zeilensumme», (2) = «Restsumme», (3) = «Totalsumme» Es sind nun (1) und (2) zwei unabhängige Schätzungen von v^2, beide verteilt wie χ^2 mit den entsprechenden Freiheitsgraden.

Der Quotient $Q = \dfrac{v^2(1)}{v^2(2)}$, also $Q = \dfrac{(N-i) \sum_i n_i (\bar{x}_i - \bar{\bar{x}})^2}{(i-1) \sum_{ik} (x_{ik} - \bar{x}_i)^2}$

ist deshalb verteilt wie f, mit $i-1$ bzw. $N-i$ Freiheitsgraden. Damit aber die Annahme, dass alle Grundgesamtheiten das gleiche Mittel aufweisen, aufrechterhalten werden kann, dürfen die beiden Streuungen (1) und (2) nur in zufälliger Weise voneinander abweichen. Ist das der Fall, so haben wir keinen Grund, an der Homogenität der Gruppe zu zweifeln, d. h., wir können wirklich annehmen, dass alle Stichprobenreihen derselben Grundgesamtheit entstammen.

Für die spätere Rechnung seien nun noch zwei vereinfachende Identitäten erwähnt:

$$\sum_{ik} (x_{ik} - \bar{\bar{x}})^2 = \sum_{ik} x_{ik}^2 - N \bar{\bar{x}}^2$$

und

$$\sum_i n_i (\bar{x}_i - \bar{\bar{x}})^2 = \sum_i (n_i \bar{x}_i) \bar{x}_i - N \bar{\bar{x}}^2.$$

III. Abschnitt

Untersuchung von Stichproben auf Normalverteilung

Dieser Abschnitt ist der Abklärung verschiedener Fragen gewidmet; erstens wollen wir prüfen, aus wievielen Elementen n eine Stichprobe zusammengesetzt sein muss, um Normalverteilung zu gewährleisten, und ob sich m, der Mittelwert der Stichproben, nur durch Zufall von μ, dem Mittel der Grundgesamtheit, unterscheidet; zweitens müssen wir die Grösse der Streuung s^2 der $\bar{c}$ aus verschiedenen Stichproben empirisch festlegen, und drittens wollen wir untersuchen, ob sich die Streuung s^2 mit zunehmender Zahl der Elemente n wirklich im Verhältnis $\frac{1}{n}$ verkleinert. Die zweite Frage überrascht vielleicht etwas, denn man würde doch bestimmt eher erwarten, dass gefragt wird, ob sich die Streuung s^2 nur durch Zufall von der durch n dividierten Streuung σ^2 der Grundgesamtheit unterscheide. — Wir berechnen

$$\sigma^2 = \frac{\sum_{i=1}^{k} S_i \,(c^{s\,i-50} - \bar{c})^2}{\sum_{i=1}^{k} S_i} = \frac{\sum_{i=1}^{k} S_i \,(c^{s\,i-50})^2}{\sum_{i=1}^{k} S_i} - \bar{c}^2,$$

wobei wir annehmen, wir hätten k verschiedene vorkommende Schlussalter. Es ist dabei aber zu beachten, dass mit dieser Streuung die Gliederung der Versicherungssummen S innerhalb eines bestimmten Schlussalters nicht erfasst werden kann, da für jedes vorkommende c^{s-50} über sämtliche S summiert wird. Für die Streuung s^2 der $\bar{c}$ hat aber gerade die Gliederung der S grosse Bedeutung, denn für ein einzelnes Element einer Stichprobe kann man zu einem der k verschiedenen Schlussalter ganz unterschiedlich grosse Versicherungssummen ziehen. Wir werden deshalb vermuten, dass s^2 grösser ausfällt als $\frac{\sigma^2}{n}$ mit einem so deutlichen Unterschied, dass er kaum mehr als zufällige Abweichung erklärt werden kann. So bleibt uns nichts anderes übrig, als s^2 aus den Versuchen zu bestimmen und uns von Fall zu Fall zu fragen, ob wir ein bestimmtes s^2 auch wirklich weiter verwenden dürfen.

Um die vorstehend aufgeführten Fragen beantworten zu können, haben wir 17 000 im Jahre 1948 abgelaufene gemischte Versicherungen des Volksversicherungsbestandes der «Basler Lebens-Versicherungs-Gesellschaft» als Grundgesamtheit bezeichnet. Der Untersuchungsbestand wies Versicherungsdauern von 10 bis 22 Jahren auf, wobei im Tarif für Erwachsene nur Dauern von 10, 15, 20 Jahre vorgesehen waren. Im gleichen Bestande war aber auch eine sehr grosse Anzahl Kinderversicherungen (ca. 6000) enthalten, so dass wir trotz des Fehlens langer Dauern einen ordentlich umfassenden Bestand vor uns hatten. Den Berechnungen legen wir die Schweizerische Volkssterbetafel SM 1939/44, ausgeglichen nach *Makeham* für Alter $\geqq 36$ mit $c = 1{,}0985\ 1878$, zugrunde. (Diese Tafel wurde uns freundlicherweise vom Eidgenössischen Statistischen Amt zur Verfügung gestellt.) Der technische Zinsfuss beträgt $2\frac{3}{4}\%$. Im übrigen sei bei dieser Gelegenheit darauf hingewiesen, dass es nicht zur Aufgabe dieser Arbeit gehört, auch Aussagen über die Genauigkeit des Verfahrens von *Lidstone* im Vergleich zur Einzelrechnung zu machen.

Wir haben vorerst unter vollständiger Ausschöpfung des Bestandes Stichproben zu 50, 100 und 200 Versicherungen gebildet. Mit diesem Vorgehen sind die einzelnen Stichproben nicht mehr ganz unabhängig voneinander, aber das Ergebnis wird durch diese Art des Ziehens praktisch kaum beeinflusst werden. Anschliessend haben wir auch grössere Stichproben, nämlich zu 300, 400, 600 und 1000, gebildet, indem wir eine entsprechende Anzahl Stichproben zu 100 Karten zusammenfassten, dabei aber die Versicherungssumme als Gewicht nicht mehr berücksichtigten. Die Berechtigung zu diesem Vorgehen werden wir mit Hilfe einer Korrelationsuntersuchung für Stichproben zu 100 Karten zeigen. Zum Schluss haben wir aus unserem Untersuchungsbestande einen Teilbestand von 12 100 Versicherungen herausgezogen, indem wir uns auf Verträge mit Versicherungsdauern $\leqq 15$ Jahren beschränkten; später wählten wir einen Bestand von 5100 Versicherungen mit Versicherungsdauern $\geqq 20$ Jahren und haben daraus je Stichproben zu 100 Karten gebildet.

a) Stichproben zu 50 Karten

Für dieses Beispiel wollen wir den Gang der Rechnungen voll angeben.

Das $\bar{c}$ des Untersuchungsbestandes liegt bei $0{,}9145 = \mu$, was einem durchschnittlichen Schlussalter von 49,05 Jahren entspricht.

Klasseneinteilung für $\bar{c}$	Intervall x	Beob. Häufigkeit $f(x)$	$x f(x)$	$x^2 f(x)$	$x^3 f(x)$
0,190–0,354	—4	1	— 4	16	— 64
0,355–0,519	—3	5	— 15	45	—135
0,520–0,684	—2	52	—104	208	—416
0,685–0,849	—1	88	— 88	88	— 88
0,850–1,014	0	82	—211	0	—703
1,015–1,179	1	71	71	71	71
1,180–1,344	2	29	58	116	232
1,345–1,509	3	9	27	81	243
1,510–1,674	4	3	12	48	192
		$N = 340$	168	673	738

$$m_1 = \frac{-43}{340} = -0{,}1265, \qquad m_2 = \frac{673}{340} - m_1^2 = 1{,}9634,$$

$$m_3 = \frac{35}{340} - 3 m_1 m_2 - m_1^3 = 0{,}8480.$$

Der Mittelwert der Stichproben mit Klasseneinteilung liegt bei $m = 0{,}9111$; $\sigma^2 = 1{,}771$. Die Streuung s_{kl}^2 einer in Klassen angeordneten Verteilung berechnen wir unter Verwendung der *Sheppardschen* Korrektur [13] für die Klasseneinteilung zu

$$s_{kl}^2 = \frac{1}{N-1} \sum x^2 f(x) - m_1^2 - 0{,}0833.$$

Es wird

$$s_{kl}^2 = \frac{673}{339} - m_1^2 - 0{,}0833 = 1{,}8859, \quad s_{kl} = 1{,}3733;$$

$$s^2 = (0{,}165)^2 s_{kl}^2 = 0{,}05134, \quad \frac{\sigma^2}{n} = \frac{1{,}771}{50} = 0{,}03542.$$

Die γ-Prüfung auf Normalverteilung gibt folgende Werte:

$k_2 = 1{,}9692,$

$k_3 = 0{,}8555,$

$$\gamma = \sqrt{\frac{(N-1)(N-2)}{6N}} \frac{k_3}{k_2 \sqrt{k_2}} = 2{,}32.$$

γ liegt ausserhalb der 5%-Schwelle von 1,96. Die Wahrscheinlichkeit, dass die vorliegende Verteilung noch durch Zufall aus einer Normalverteilung entstanden sein könnte, ist $P = 2\%$.

Die χ^2-Prüfung gibt folgende Werte

Untere Intervallgrenze x	$x - m_1$	$\xi = \frac{x-m_1}{s_{kl}}$ Standardisiert	$\int_{-\infty}^{\xi} dF$	Relative Häufigkeit	Theoretische Häufigkeit	Beobachtete Häufigkeit	χ^2
				0,000 74	0,25		
— 4,5	— 4,3735	— 3,18	0,000 74	0,000 21	2,11	1	0,58
— 3,5	— 3,3735	— 2,46	0,006 95	0,034 86	11,85	5	3,96
— 2,5	— 2,3735	— 1,73	0,041 81	0,116 84	39,72	52	3,80
— 1,5	— 1,3735	— 1,00	0,158 65	0,234 17	79,62	88	0,88
— 0,5	— 0,3735	— 0,272	0,392 82	0,282 98	96,21	82	2,10
0,5	0,6265	0,456	0,675 80	0,205 20	69,77	71	0,02
1,5	1,6265	1,18	0,881 00	0,090 93	30,92	29	0,12
2,5	2,6265	1,91	0,971 93	0,023 93	8,14	9	0,09
3,5	3,6265	2,64	0,995 86	0,003 76	1,28	3	2,31
4,5	4,6265	3,37	0,999 62	0,000 38	0,13		
				1,000 00	340,00		13,86

Bei 6 Freiheitsgraden beträgt $\chi^2 = 13{,}86$. Die Wahrscheinlichkeit dafür, dass die vorliegende Verteilung noch durch Zufall aus einer Normalverteilung entstanden sein könnte, ist $P = 3{,}5\%$.

Mit Hilfe der *t-Verteilung* haben wir die Möglichkeit, die Abweichung zwischen m, dem Mittel aus den Stichproben, und μ, dem Mittel der Grundgesamtheit, zu prüfen. Es ist

$$t = \frac{(\mu - m)}{s}\sqrt{(N-1)} = \frac{0{,}9145 - 0{,}9111}{0{,}2266}\sqrt{339} = 0{,}276.$$

Somit liegt t weit innerhalb der 5%-Schwelle von 1,967. Ebenso sind wir in der Lage, den Unterschied zwischen der mit Stichproben zu 50 Karten erwarteten Streuung $\sigma^2_{\frac{1}{50}} = 0{,}03542$ und s^2 zu prüfen. Es ist

$$\xi^2 = \frac{s^2(N-1)}{\sigma^2_{\frac{1}{50}}} = \frac{0{,}05134 \cdot 339}{0{,}03542} = 491{,}4$$

verteilt wie χ^2 mit $(N-1) = n'$ Freiheitsgraden. Der Unterschied $x = \sqrt{2\xi^2} - \sqrt{2(n'-1)} = \sqrt{982{,}8} - \sqrt{676} = 5{,}35$ ist normal verteilt und liegt weit ausserhalb der 5% Sicherheitsschwelle von 1,960. Wir erhalten das erwartete Resultat, dass die beobachtete Abweichung zwischen s^2 und $\frac{\sigma^2}{50}$ nur etwa mit einer Wahrscheinlichkeit von 10^{-7} noch durch Zufall entstanden sein könnte.

Zusammenfassend dürfen wir feststellen, dass die Mittelwerte aus Stichproben zu 50 Karten kaum als normal verteilt angesprochen werden dürfen.

b) Stichproben zu 100 Karten

In den folgenden Tabellen geben wir abgekürzt die Verteilung der Mittelwerte sowie die Wahrscheinlichkeiten zu den Prüfverfahren.

Klassen-einteilung	Beob-achtete Häufig-keit	Theoretische Häufigkeit	
		0,11	
0,405–0,549	1	1,71	$\gamma = 1{,}188$, $P = 23\%$
0,550–0,694	10	12,42	
0,695–0,839	49	40,14	$\chi^2 = 4{,}32$ mit 4 Freiheitsgraden,
0,840–0,984	56	60,50	$P = 37\,\%$
0,985–1,129	36	40,34	$m = 0{,}9137$
1,130–1,274	16	12,86	$s^2 = 0{,}02512$
1,275–1,419	2	1,80	$\frac{\sigma^2}{100} = 0{,}01771$ } $P \approx 0{,}05\,\%$
		0,12	
	170	170,00	

Die Wahrscheinlichkeiten für die Abweichungen zwischen m und μ werden wir für sämtliche nun folgenden Beispiele zusammen am Schlusse dieses Abschnittes in einer Tabelle darstellen.

Neben s^2, das aus der vorliegenden Verteilung durch Klassenbildung gewonnen wurde und das auch hier wieder nur mit einer ganz geringen Wahrscheinlichkeit durch Zufall aus $\frac{\sigma^2}{100}$ entstanden sein könnte, haben wir aus den 170 Stichproben auch das genaue $\bar{s}^2 = 0{,}02470$ durch Einzelrechnung gebildet. Die Wahrscheinlichkeit, dass s^2 durch Zufall aus $\bar{s}^2$ entstanden sein könnte, hat, bei einem x von 0,21, den Wert $P = 83\,\%$. Wir müssen aber auch eine Wahrscheinlichkeit

in dieser Höhe erhalten, denn die Differenz zwischen den beiden Streuungen ist nur wegen der Klassenbildung entstanden. Da unser $\bar{s}^2$ schon eine ziemlich zuverlässige Schätzung für die wirkliche Streuung der Verteilung der Mittelwerte in Stichproben zu 100 Karten ist, wollen wir es in Zukunft auch als diese wirkliche Streuung betrachten. Für die Prüfung der Wahrscheinlichkeit, dass in n-mal grösseren Stichproben die Streuung im Verhältnis $\frac{1}{n}$ kleiner wird, wollen wir deshalb $\bar{s}^2$ als theoretische Streuung verwenden. Die Resultate für diese Prüfung sind auch hier wieder in einer Tabelle am Schlusse des Abschnittes zusammengestellt.

c) Stichproben zu 200 Karten

I = 200 Karten mit Versicherungssumme als Gewicht.
II = 2 × 100 Karten ohne Gewicht.

Klasseneinteilung	Beobachtete Häufigkeit		Theoretische Häufigkeit	
	I	II	I	II
			0,15	0,16
0,600–0,689	1	1	1,40	1,47
0,690–0,779	9	9	7,27	7,35
0,780–0,869	18	19	19,44	19,38
0,870–0,959	28	26	27,26	27,07
0,960–1,049	18	19	20,02	19,95
1,050–1,139	10	10	7,74	7,82
1,140–1,229	1	1	1,54	1,61
			0,18	0,19
	85	85	85,00	85,00

Bei beiden Verteilungen liegt der Mittelwert m bei $m = 0{,}9166$.

Für die beiden Streuungen s^2 erhalten wir: (I) $s^2 = 0{,}01186$,
(II) $s^2 = 0{,}01205$;

Für die beiden γ finden wir: $\gamma_{\mathrm{I}} = -0{,}0272$, $P = 98\,\%$,
$\gamma_{\mathrm{II}} = -0{,}0298$, $P = 98\,\%$.

Schliesslich bekommen wir für die beiden χ^2 bei je 4 Freiheitsgraden:
$\chi^2_{\mathrm{I}} = 1{,}70$, $P = 79\,\%$; $\chi^2_{\mathrm{II}} = 1{,}46$, $P = 83\,\%$.

Wir sehen, dass schon für Stichproben zu 200 Karten die Normalverteilung in einem Masse gewährleistet ist, wie sie auch für grössere Stichproben kaum mehr übertroffen werden dürfte. Auch gibt uns der praktische Versuch das Anrecht, in grösseren Stichproben die Summen nicht mehr zu berücksichtigen. Zur genaueren Begründung wollen wir die Korrelation zwischen der Steigung der $\bar{c}$ und der Zunahme der Versicherungssumme in Stichproben zu 100 Karten betrachten. Sind die beiden Reihen korrelationslos, so können wir mit gutem Recht bei grösseren Stichproben die Versicherungssumme, die bei der Bildung der $\bar{c}$ als Gewicht dient, vernachlässigen. Das will heissen, dass wir beispielsweise für die Bildung einer Stichprobe zu 200 Karten einfach 2 Stichproben zu 100 Karten zusammenlegen können und als $\bar{c}$ das gewöhnliche arithmetische Mittel der beiden $\bar{c}$ der Stichproben zu 100 Karten nehmen.

Korrelation zwischen Versicherungssumme und $\bar{c}$ in Stichproben zu 100 Karten

Klassenmitten der Versicherungssummen	Klassenmitten der $\bar{c}$								
	0,515	0,615	0,715	0,815	0,915	1,015	1,115	1,215	1,315
88 000	.	.	1	.	.	.	.	.	.
92 000	.	.	2	1	1	2	1	1	.
96 000	.	1	2	10	3	5	1	2	.
100 000	.	2	6	8	8	7	4	.	.
104 000	1	3	3	9	11	6	5	2	1
108 000	.	1	3	9	9	6	2	1	.
112 000	.	.	2	6	3	2	4	1	2
116 000	.	.	.	1	2	1	2	.	.
120 000	.	1	1	.	.	.	.	.	.
124 000	.	.	.	.	.	.	.	1	.
128 000	.	.	.	.	.	.	1	.	.

Der Korrelationskoeffizient r wird in diesem Falle berechnet zu $r = +0{,}124$, so dass wir tatsächlich von vollständiger Unabhängigkeit sprechen können. Für das Bestimmtheitsmass B ergibt sich $B = 0{,}0154$, d. h., die Veränderung des $\bar{c}$ lässt sich nur zu 1,54 % aus der Zunahme der Versicherungssumme erklären.

d) Stichproben zu 3 × 100, 4 × 100, 6 × 100, 10 × 100 Karten

Klassen-einteilungen	Beob-achtete Häufig-keit	Theoretische Häufigkeit	
3 × 100		0,20	
0,660–0,759	6	3,90	$\gamma = -0{,}554$, $P = 58\,\%$
0,760–0,859	18	23,25	
0,860–0,959	48	43,01	$\chi^2 = 3{,}10$ bei 2 Freiheitsgraden, $P = 21\,\%$
0,960–1,059	23	24,89	$m = 0{,}9125$
1,060–1,159	5	4,50	
		0,25	$s^2 = 0{,}00774$
	100	100,00	
4 × 100		0,19	
0,695–0,784	5	3,90	$\gamma = -0{,}404$, $P = 68\,\%$
0,785–0,874	22	23,77	
0,875–0,964	44	43,50	$\chi^2 = 0{,}47$ bei 2 Freiheitsgraden, $P = 79\,\%$
0,965–1,054	25	24,32	$m = 0{,}9204$
1,055–1,144	4	4,11	
		0,21	$s^2 = 0{,}006\,116$
	100	100,00	
6 × 100		0,32	
0,740–0,809	5	4,83	$\gamma = 0{,}0814$, $P = 93\,\%$
0,810–0,879	26	24,48	
0,880–0,949	39	41,59	$\chi^2 = 0{,}35$ bei 2 Freiheitsgraden, $P = 84\,\%$
0,950–1,019	25	23,83	$m = 0{,}9138$
1,020–1,089	5	4,64	
		0,31	$s^2 = 0{,}004\,095$
	100	100,00	
10 × 100		0,03	
0,745–0,809	1	1,37	$\gamma = -0{,}0954$, $P = 92\,\%$
0,810–0,874	15	13,50	
0,875–0,939	27	29,34	$\chi^2 = 0{,}85$ bei 2 Freiheitsgraden, $P = 65\,\%$
0,940–1,004	16	14,23	$m = 0{,}9081$
1,005–1,069	1	1,50	
		0,03	$s^2 = 0{,}002\,44$
	60	60,00	

e) Stichproben zu 100 Karten mit Versicherungsdauern ⩾ 15 Jahre

Klassen-einteilung	Beob-achtete Häufig-keit	Theoretische Häufigkeit	
		0,62	
0,430–0,544	3	3,45	$\bar{c}$ Untersuchungsbestand = 0,8325, entspricht 48,05 Jahren
0,545–0,659	13	12,35	
0,660–0,774	31	27,07	
0,775–0,889	32	34,92	$\gamma = 0{,}897$, $P = 37\,\%$
0,890–1,004	26	26,69	$\chi^2 = 3{,}64$ bei 4 Freiheitsgraden, $P = 46\,\%$
1,005–1,119	10	12,09	
1,120–1,234	6	3,24	$m = 0{,}8301$, $s^2 = 0{,}024\,24$
		0,57	
	121	121,00	$\bar{s}^2 = 0{,}024\,56$

f) Stichproben zu 100 Karten mit Versicherungsdauern ⩾ 20 Jahre

Klassen-einteilung	Beob-achtete Häufig-keit	Theoretische Häufigkeit	
		0,07	
0,350–0,459	0	3,76	$\bar{c}$ Untersuchungsbestand = 0,8452, entspricht 48,20 Jahren
0,460–0,569	1		
0,570–0,679	16	10,57	
0,680–0,789	19	21,88	$\gamma = 1{,}665$, $P = 10\,\%$
0,790–0,899	32	28,29	$\chi^2 = 6{,}61$ bei 4 Freiheitsgraden, $P = 16\,\%$
0,900–1,009	19	22,45	
1,010–1,119	10	10,98	$m = 0{,}8467$
1,120–1,229	3	3,92	$s^2 = 0{,}024\,15$
1,230–1,339	2		
		0,08	$\bar{s}^2 = 0{,}024\,005$
	102	102,00	

Wir haben unter *e* und *f* keine grösseren Stichproben erhoben, weil wir annehmen dürfen, dass auch grössere Stichproben normal verteilt sind.

Zusammenfassend können wir sagen, dass schon Stichproben zu 100 Karten mit guter Wahrscheinlichkeit normal verteilt sind, dass

die m der Verteilungen sich nur durch Zufall von den $\bar{c} = \mu$ der Grundgesamtheiten unterscheiden (wie die folgende Tabelle zeigt), dass die Streuungen s^2 grosser Stichproben im Verhältnis $\frac{1}{n}$ abnehmen und sich nur durch Zufall aus einer einmal angenommenen Vergleichsstreuung $\bar{s}^2$ entstanden denken lassen (wie die übernächste Tabelle zeigt).

Wahrscheinlichkeiten für die Abweichung zwischen m und μ

N	n	μ	m	$t = \frac{(m-\mu)}{s}\sqrt{(N-1)}$ $t \approx$ normal verteilt für $N > 30$	Wahrscheinlichkeit des zufälligen Abweichens %
340	50	0,9145	0,9111	0,276	78
170	100	0,9145	0,9137	0,066	94
85	200	0,9145	0,9166	0,176	86
85	2 × 100	0,9145	0,9166	0,175	86
100	3 × 100	0,9145	0,9125	0,226	82
100	4 × 100	0,9145	0,9204	0,751	45
100	6 × 100	0,9145	0,9138	0,109	91
60	10 × 100	0,9145	0,9081	0,995	32
121	100	0,8325	0,8301	0,169	86
102	100	0,8452	0,8467	0,097	92

Wahrscheinlichkeiten für die Abweichung zwischen s^2 und $\bar{s}^2$

N	n	$\bar{s}^2$	s^2	$x = \sqrt{\frac{2(N-1)s^2}{\bar{s}^2}} - \sqrt{2(N-1)-2}$ $x \approx$ normal verteilt für $N > 30$	Wahrscheinlichkeit des zufälligen Abweichens %
340	50	0,049 4	0,051 34	0,54	58
170	100	0,024 7	0,025 12	0,21	83
85	200	0,012 35	0,011 86	— 0,18	85
85	2 × 100	0,012 35	0,012 05	— 0,08	93
100	3 × 100	0,008 233	0,007 74	— 0,36	72
100	4 × 100	0,006 175	0,006 116	0,01	99
100	6 × 100	0,004 117	0,004 095	0,03	97
60	10 × 100	0,002 47	0,002 44	0,03	97
121	100	0,024 56	0,024 24	— 0,04	97
102	100	0,024 005	0,024 15	0,12	90

IV. Abschnitt

§ 1. Durchschnittliche Schlussalter und Reserven eines effektiven Versicherungsbestandes

Bei der Prüfung auf Normalverteilung im vorhergehenden Abschnitt haben wir immer nur die Versicherungen eines einzigen Ablaufjahres betrachtet. Die nun folgenden Untersuchungen werden sich auf effektive Bestände stützen, d. h. auf Bestände, bei denen Versicherungen über alle möglichen Ablaufjahre vorkommen. Zu diesem Zwecke haben wir aus dem Volksversicherungsbestande der «Basler Lebens-Versicherungs-Gesellschaft» die Versicherungen mit Beginn in der 52. Woche herausgegriffen. Diese Auswahl, welche Ablaufjahre von 1950 bis 1983 enthielt, stellt an und für sich nur eine Stichprobe einer umfassenden Gesamtheit dar, kann aber mit 23 153 Versicherungen schon als sehr gross und als Grundgesamtheit dienlich angesprochen werden. Allerdings mussten wir den Bestand in zwei Teile aufspalten, weil ursprünglich die Kinderversicherungen auf das Leben des Kindes allein abgeschlossen wurden, während sie heute als sog. kombinierte Versicherungen sowohl auf das Leben des Kindes als auch des Prämienzahlers laufen und die Reserverechnung sich nur nach dem Alter des Prämienzahlers richtet. Wir haben aufgeteilt in einen Bestand I, enthaltend 11 496 bis April 1944 abgeschlossene Versicherungen, und in einen Bestand II, enthaltend 11 657 seit April 1944 abgeschlossene Versicherungen.

In der folgenden Tabelle geben wir eine Übersicht über die Gliederung dieser beiden Untersuchungsbestände, wobei wir für jedes einzelne Ablaufjahr das durchschnittliche Schlussalter und die *Lidstonereserve* auf die Jahresmitte 1949 angegeben haben. Auch sei bei dieser Gelegenheit erwähnt, dass sämtliche später erwähnten Reserven auf den gleichen Zeitpunkt berechnet sind.

Genaue Schlussalter und Reserven des Bestandes der 52. Woche

Bestand I				Bestand II			
Gruppe	Ablaufjahr	Schlussalter	Reserve	Gruppe	Ablaufjahr	Schlussalter	Reserve
1	1950	49,80	578 743	1	1954	55,35	105 258
	1951	48,95	763 212		1955	54,80	84 577
	1952	49,30	896 568		1956	55,50	82 537
	1953	49,40	636 218		1957	57,20	69 462
2	1954	46,25	337 150	2	1958	54,55	164 374
	1955	48,80	498 364		1959	54,10	231 069
	1956	49,45	689 603		1960	54,40	139 470
	1957	49,70	733 461		1961	56,15	106 987
					1962	55,75	106 754
3	1958	47,85	526 315	3	1963	50,60	306 912
	1959	49,85	222 578		1964	51,80	386 793
	1960	49,85	330 254		1965	53,85	229 634
	1961	49,80	451 355		1966	52,80	211 516
	1962	49,55	551 895		1967	52,00	163 688
					1968	52,15	234 129
4	1963	52,40	237 471	4	1969	54,60	203 359
	1964	51,90	87 644		1970	54,75	111 622
	1965	51,55	114 761		1971	54,85	92 764
	1966	52,70	183 492		1972	54,15	72 934
	1967	52,75	239 444		1973	53,95	75 493
5	1968	53,20	57 992	5	1974	55,10	62 909
	1969	51,65	10 223		1975	54,50	38 863
	1970	51,70	17 099		1976	56,45	29 219
	1971	56,00	52 241		1977	56,15	22 648
	1972	53,75	62 709		1978	55,90	19 571
6	1973	58,15	7 607	6	1979	58,10	12 714
	1974	52,75	1 249		1980	58,50	5 985
	1975	51,40	669		1981	56,95	3 235
	1976	58,50	7 657		1982	56,30	754
	1977	57,60	12 311		1983	57,60	— 630
			8 308 285				3 374 600

§ 2. Untersuchung auf Homogenität

In der letzten Tabelle fällt auf, dass eine Zusammenfassung in Gruppen von je 4–6 Jahren vorgenommen worden ist, derart, dass in einer Gruppe jeweils ungefähr das gleiche Schlussalter auftritt. Dies legt die Vermutung nahe, dass alle Ablaufjahre einer Gruppe ein übereinstimmendes Schlussalter aufweisen könnten, oder mit anderen Worten, dass die Gruppe in bezug auf das Schlussalter homogen ist und die Abweichungen nur durch Zufall entstanden sind. Zu diesem Ergebnis kommen wir vor allem durch die Annahme, das Schlussalter hänge sehr stark von der Versicherungsdauer ab. Da in der Regel die durch 5 ohne Rest teilbaren Dauern stark überwiegen, so wird man erwarten, dass beispielsweise in dem Ablaufjahr mit den Dauern von 10 bis 35 ein anderes Schlussalter vorkommt als bei einem Ablaufjahr mit Dauern von 11 bis 35. Anderseits wird aber das Jahr mit Dauern von 11 bis 35 ziemlich genau das gleiche Schlussalter aufweisen wie das Jahr mit Dauern von 15 bis 35, denn die Dauern von 11 bis 14 bleiben durch ihre geringe Anzahl sozusagen ohne Einfluss. Im weiteren möge man sich durch die nicht ganz regelmässige Gruppeneinteilung nicht stören lassen, denn sie ist durch den Umstand bedingt, dass sich in einzelnen Jahren zwei 5er Dauern überschneiden, so dass wir diese Jahre der Gruppe zuordnen mussten, bei der eine der beiden Dauern mehr Gewicht hat.

Um zu entscheiden, ob die Abweichungen der Schlussalter in einer Gruppe wirklich nur durch Zufall entstanden sind, stützen wir uns auf das früher angeführte Verfahren zur Prüfung der Homogenität. Bilden wir in jedem Ablaufjahr mehrere Stichproben zu 100 Karten, so sind wir in der Lage, die verlangten drei Streuungen zu berechnen. Es ist dies einmal die Streuung zwischen den Mitteln der einzelnen Ablaufjahre und dem Mittel der Gruppe, die wir «Zeilenstreuung» nennen wollen, dann die Summe der quadratischen Abweichungen der Stichproben eines einzelnen Ablaufjahres um ihr Mittel, summiert über sämtliche Jahre einer Gruppe und dividiert durch die Gesamtzahl Stichproben weniger Anzahl Jahre einer Gruppe. Diese Streuung wollen wir «Reststreuung» nennen. Als dritte mögliche Streuung bleibt uns noch diejenige, die wir schon rein gefühlsmässig als die eigentliche Streuung der Stichproben in einer Gruppe bezeichnen werden, nämlich die Streuung zwischen sämtlichen Stichproben einer Gruppe und dem Mittel der Gruppe, die wir «Totalstreuung» nennen wollen.

Für das Verfahren können wir zwei Fälle unterscheiden. Ist die «Zeilenstreuung» kleiner als die «Reststreuung», dann haben wir keinen Grund, an der Homogenität der Gruppe zu zweifeln, denn dann haben die Mittel der einzelnen Ablaufjahre eine kleinere Streuung als die Stichproben aus den verschiedenen Ablaufjahren. Für den umgekehrten Fall können wir Wahrscheinlichkeitsgrenzen setzen, bei denen die beiden Streuungen und damit auch die Schlussalter der verschiedenen Ablaufjahre noch durch Zufall voneinander abweichen. Der Quotient $\frac{\text{«Zeilenstreuung»}}{\text{«Reststreuung»}}$ ist dabei verteilt wie die Masszahl f.

Setzen wir als Grenze eine Wahrscheinlichkeit von 5 % und liegt die Abweichung unserer beiden Streuungen noch innerhalb dieser Grenze, dann haben wir auch hier keinen Grund, an der Homogenität der Gruppe zu zweifeln.

Prüfung auf Homogenität

Bestand I				
Summe		Freiheitsgrade	Streuung	
1950–1953				
Zeilensumme	0,001 22	3	0,000 407	Quotient
Restsumme	0,675 98	26	0,026 00	$= 0{,}0157 < 1$
Totalsumme	0,677 20	29	0,023 35	5% $f = 2{,}98$
1954–1957				
Zeilensumme	0,2232	3	0,074 40	$f = 3{,}22$
Restsumme	0,5767	25	0,023 07	5 % $f = 2{,}99$
Totalsumme	0,7999	28	0,028 57	1 % $f = 4{,}68$
1958–1962				
Zeilensumme	0,1398	4	0,034 95	$f = 2{,}48$
Restsumme	0,3811	27	0,014 11	5 % $f = 2{,}73$
Totalsumme	0,5209	31	0,016 80	
1963–1967				
Zeilensumme	0,0524	4	0,013 10	Quotient $= 0{,}95 < 1$
Restsumme	0,1241	9	0,013 79	5 % $f = 3{,}63$
Totalsumme	0,1765	13	0,013 58	

Bestand II				
Summe		Freiheitsgrade	Streuung	
1954–1957				
Zeilensumme	0,1705	3	0,056 83	Quotient = 0,83 <1
Restsumme	0,2743	4	0,068 58	5 % $f = 6{,}59$
Totalsumme	0,4448	7	0,063 54	
1958–1962				
Zeilensumme	0,4619	4	0,115 48	$f = 2{,}87$
Restsumme	0,6438	16	0,040 24	5 % $f = 3{,}01$
Totalsumme	1,1057	20	0,055 29	
1963–1968				
Zeilensumme	0,6797	5	0,135 94	$f = 9{,}69$
Restsumme	0,6735	48	0,014 03	5 % $f = 2{,}4$
Totalsumme	1,3532	53	0,025 53	1 ‰ $f = 4{,}9$
1969–1973				
Zeilensumme	0,0480	4	0,012 00	$f = 1{,}46$
Restsumme	0,0987	12	0,008 23	5 % $f = 3{,}26$
Totalsumme	0,1467	16	0,009 17	

Als Ergebnis des Versuches können wir festhalten, dass bei sechs Gruppen kaum ein Zweifel über die Homogenität möglich ist. Bei einer Gruppe liegt f ganz knapp ausserhalb der 5 %-Grenze, und nur in einem Falle verneint das Verfahren die Richtigkeit der Annahme, was aber durch eine abnormale Störung durch Kinderversicherungen erklärt werden kann.

Gruppenschlussalter für den Bestand der 52. Woche

Bestand I		Bestand II	
Gruppe	Schlussalter	Gruppe	Schlussalter
1	49,35	1	55,80
2	49,05	2	54,95
3	49,35	3	52,15
4	52,45	4	54,45
5	54,05	5	55,55
6	57,70	6	57,70

Für diese Gruppenalter folgt für die Reserve auf Mitte 1949 bei Bestand I der Betrag 8 307 158, was einer Abweichung von 0,0136 % von der *Lidstonereserve* entspricht; für Bestand II ist der Wert 3 372 857, mit einer Abweichung von 0,0517 %. Auch die Reserverechnung bestätigt uns somit unsere Annahme.

§ 3. Konstruktion eines theoretischen Bestandes

Betrachten wir die Schlussalter bei sämtlichen Ablaufjahren, die zu einer bestimmten Dauer gehören, dann sehen wir, dass diese Schlussalter nur sehr wenig voneinander abweichen. Es sei hier als einziges Beispiel die Dauer 20 des Bestandes I herausgegriffen.

Schlussalter für Versicherungsdauern 20 Jahre Bestand I

Jahr	Schlussalter	Versicherungssumme	Jahr	Schlussalter	Versicherungssumme
1953	47,25	254 185	1958	48,15	801 795
1954	46,45	291 950	1959	51,35	218 365
1955	49,25	335 480	1960	48,60	490 420
1956	47,45	583 300	1961	48,40	694 250
1957	48,40	672 120	1962	47,60	1 011 525

Es ist überdies ersichtlich, dass die grössten Abweichungen gerade bei den kleinsten Versicherungssummen auftreten. Dies legt uns die Annahme nahe, dass bei einem sehr grossen Bestand das Schlussalter für eine bestimmte Versicherungsdauer gegen den Mittelwert der Schlussalter aller Ablaufjahre für die gleiche Dauer geht. Gestützt darauf haben wir einen «theoretischen Bestand» konstruiert, bei dem wir in jedem einzelnen Ablaufjahr für jede vorkommende Versicherungsdauer das entsprechende gemeinsame Schlussalter der gleichen Versicherungsdauer aus dem ganzen Bestande mit der Versicherungssumme als Gewicht eingesetzt haben. Daraus haben wir dann das theoretische Schlussalter eines einzelnen Ablaufjahres erhalten, wobei wir glauben, dass wir derart einen Bestand gebildet haben, der im Aufbau einem umfassenden Bestand ziemlich nahe kommt und dass wir ihn in seiner späteren Bedeutung als Bezugspunkt für Wahrscheinlichkeiten von Reserveabweichungen benützen können.

Schlussalter und Reserven für die theoretischen Bestände

Theoretischer Bestand I			Theoretischer Bestand II		
Gruppe	Schlussalter	Schlussalter der Gruppe	Gruppe	Schlussalter	Schlussalter der Gruppe
1	49,55	49,85	1	54,95	56,05
	50,10			56,05	
	49,90			56,30	
	49,80			56,60	
2	47,55	49,20	2	54,95	54,85
	49,25			54,60	
	49,45			54,90	
	49,50			55,00	
				54,95	
3	48,20	49,25	3	50,90	52,40
	49,05			51,90	
	49,10			53,20	
	49,60			53,10	
	49,80			52,95	
				52,45	
4	51,85	52,15	4	54,00	54,15
	51,80			54,55	
	51,80			54,25	
	52,35			54,10	
	52,40			53,95	
5	54,15	54,20	5	55,35	55,65
	54,30			55,50	
	54,20			55,60	
	54,25			55,85	
	54,20			56,00	
6	57,15	57,30	6	57,30	57,40
	57,30			57,30	
	57,30			57,55	
	57,30			57,40	
	57,30			57,65	
Reserven:	8 306 507	8 305 647	Reserven:	3 373 755	3 373 811
Abweichungen:			Abweichungen:		
	— 0,0214 %	— 0,0318 %		— 0,025 %	— 0,0234 %

§ 4. Wahrscheinlichkeiten für Reserveabweichungen

Wir haben schon früher ausgeführt, dass der Hauptzweck unserer Arbeit in der Angabe von Wahrscheinlichkeiten besteht, mit denen eine bestimmte Reserveabweichung nicht überschritten wird. Dazu wären wir bereits jetzt, gestützt auf die Bestände I und II, ohne die Gruppenbildung vorzunehmen, in der Lage, denn wir wissen, dass schon Stichproben zu 100 Karten aus einem Ablaufjahr normal verteilt sind. Die Streuung der Stichproben würde durch Versuche festzulegen sein, und dann wären Wahrscheinlichkeiten gegeben, mit denen ein bestimmtes Schlussalter als Abweichung vom genauen Schlussalter überschritten bzw. unterschritten wird. Dazu könnten wir die Reserve rechnen und hätten damit auch die Wahrscheinlichkeit der Reserveabweichung in einem einzelnen Ablaufjahr erhalten. Ein normal aufgebauter Versicherungsbestand enthält aber nun etwa 35–40 Ablaufjahre. Für die Wahrscheinlichkeit, dass sowohl in einem wie auch in allen anderen Jahren ein bestimmter Reservefehler überschritten wird, müssten wir somit das Produkt aus den 35–40 Wahrscheinlichkeiten bilden. Da aber die Abweichungen in den einzelnen Ablaufjahren ganz unterschiedliche Gewichte aufweisen, müssten wir als eine Möglichkeit die Wahrscheinlichkeitsgrenzen für jedes Jahr so klein wählen (beispielsweise 0,1), dass auch in den Jahren mit den grössten Abweichungen das zugehörige Schlussalter kaum mehr über- bzw. unterschritten würde. Die Folge davon wäre, dass wir für die Wahrscheinlichkeit einer grösseren Reserveabweichung einen dermassen kleinen Wert (etwa $0{,}1^{35}$) erhalten würden, dass wir praktisch von vollkommener Unmöglichkeit sprechen könnten. Die Reserveabweichung selbst aber würde so gross, dass das Verfahren an und für sich sinnlos gemacht wäre. Als zweite Möglichkeit könnten wir vorerst auch die Grösse der zugelassenen Reserveabweichung festhalten. Um aber noch eine sinnvolle Wahrscheinlichkeit für grössere Reserveabweichungen zu erhalten, müssten wir die Stichproben in einem einzelnen Ablaufjahr so gross wählen, dass die Rationalität des Verfahrens sehr in Frage gestellt würde.

Diesen Übelständen können wir durch die von uns vorgeschlagene Gruppenbildung weitgehend abhelfen. Bei unserer Art der Zusammenfassung sind beispielsweise 5 Stichproben zu 100 Karten, wovon je eine aus einem Ablaufjahr einer Gruppe entnommen ist, gleichwertig

einer Stichprobe zu 500 Karten, entnommen aus dem Bestande einer Gruppe, was wir früher mit der Homogenitätsprüfung gezeigt haben. Dabei erhalten wir das $\bar{c}$ der Stichprobe zu 500 Karten, indem wir das gewöhnliche arithmetische Mittel der $\bar{c}$ der Stichproben zu 100 Karten bilden. Da die Stichproben zu 500 Karten bestimmt normal verteilt sind, können wir auch für die ganze Gruppe die Wahrscheinlichkeit einer grösseren Abweichung des Schlussalters und damit der Reserve angeben. Auf diese Weise haben wir dann die Zahl der Wahrscheinlichkeitsfaktoren für den Gesamtbestand auf $\frac{1}{5}$ reduziert. Weiter benötigen wir für die Abweichung des Schlussalters einer Gruppe mit k Ablaufjahren die Streuung s^2 der Stichproben zu $k \times 100$ Karten. Wir haben aber früher gezeigt, dass diese Streuung $\frac{1}{k}$ der Streuung der Stichproben zu 100 Karten beträgt. Nun hat uns einerseits die Homogenitätsprüfung in der «Totalstreuung» bereits die Streuung der Stichproben zu 100 Karten einer Gruppe geliefert, anderseits haben wir bei den Untersuchungen auf Normalverteilung in der Streuung $\bar{s}^2$ ebenfalls die Streuung der Stichproben zu 100 Karten erhalten. Mit Hilfe der f-Verteilung wollen wir deshalb prüfen, ob die «Totalstreuungen» nur durch Zufall aus den Streuungen $\bar{s}^2$ entstanden sind.

Gruppe	«Totalstreuung»	$\bar{s}^2$	n_1	n_2	f	5 % f
Bestand I						
1	0,023 35	0,024 70	∞	29	1,058	1,64
2	0,028 57	0,024 70	28	∞	1,157	$\sim$ 1,5
3	0,016 80	0,024 56	∞	31	1,462	$\sim$ 1,6
4	0,013 58	0,024 00	∞	13	1,767	2,21
Bestand II						
1	0,063 54	0,024 70	7	∞	2,572	$\sim$ 2,0
2	0,055 29	0,024 70	20	∞	2,238	$\sim$ 1,6
3	0,025 53	0,024 56	53	∞	1,039	$\sim$ 1,35
4	0,009 17	0,024 00	∞	16	2,617	2,01

Die Prüfung zeigt, dass in Bestand II kaum von einer Übereinstimmung gesprochen werden kann, während bei Bestand I die

Abweichungen noch im Rahmen des Zufälligen bleiben. Da aber die $\bar{s}^2$ wohl auch für den Bestand I zu wenig repräsentativ sind (wir haben zur Berechnung der $\bar{s}^2$ nur Verträge mit einer maximalen Versicherungsdauer von 22 Jahren benützen können), wollen wir überall die «Totalstreuung» weiter verwenden. Für die Gruppen Nr. 5 beider Bestände setzen wir jeweils die «Zeilenstreuung» ein, da für die Berechnung der «Totalstreuung» zu wenig Elemente vorhanden sind, und für die Gruppen Nr. 6 verwenden wir die gleiche Streuung, da wir eine richtige Streuung wohl kaum mehr vernünftig berechnen könnten.

Zur Ermittlung der Reserveabweichungen wollen wir in der Folge als Wahrscheinlichkeit 20% setzen, d. h., dass höchstens in 10% aller Fälle eine grössere Abweichung als $1{,}2816 \sqrt{s}$ vom $\bar{c}$ der Gruppe nach unten und ebenfalls höchstens in 10% aller Fälle eine solche Abweichung nach oben eintritt.

Schlussalter und Reserven für 10% Wahrscheinlichkeiten

Bestand I			Theoretischer Bestand I		
Gruppe	Schlussalter		Gruppe	Schlussalter	
	untere 10%-Grenze	obere 10%-Grenze		untere 10%-Grenze	obere 10%-Grenze
1	48,15	50,40	1	48,75	50,85
2	47,70	50,25	2	47,90	50,40
3	48,50	50,15	3	48,40	50,10
4	51,85	53,00	4	51,55	52,70
5	53,00	55,00	5	53,15	55,15
6	57,00	58,40	6	56,50	58,00
Reserve:	8 293 966	8 320 415		8 292 728	8 319 171
Bestand II			Theoretischer Bestand II		
1	54,75	56,80	1	55,00	56,95
2	54,00	55,80	2	53,90	55,75
3	51,40	52,85	3	51,65	53,05
4	54,05	54,85	4	53,75	54,55
5	55,00	56,10	5	55,05	56,15
6	57,25	58,15	6	56,95	57,85
Reserve:	3 342 523	3 403 626		3 343 071	3 403 592

Der angegebenen Tabelle entnehmen wir folgende Reservefehler:

Bestand I

untere Grenze — 0,172 % von der genauen Lidstonereserve
obere Grenze + 0,146 % » » » »

Bestand II

untere Grenze — 0,951 % » » » »
obere Grenze + 0,860 % » » » »

Theoretischer Bestand I

untere Grenze — 0,166 % von der Reserve des theoretischen Best.
obere Grenze + 0,152 % » » » » » »

Theoretischer Bestand II

untere Grenze — 0,909 % » » » » » »
obere Grenze + 0,884 % » » » » » »

Gesamtbestand I und II

untere Grenze — 0,397 % von der genauen Lidstonereserve
obere Grenze + 0,352 % » » » »

Theoretischer Gesamtbestand I und II

untere Grenze — 0,381 % von der Reserve des theoretischen Best.
obere Grenze + 0,364 % » » » » » »

Für Bestand I sind die Reserveabweichungen wirklich klein, während der Bestand II als neuer Versicherungsbestand den ungünstigsten Fall darstellt, wo die grossen Reserven mit kurzer restlicher Dauer, bei denen ein Fehler im durchschnittlichen Schlussalter sozusagen ohne Einfluss bleibt, ganz fehlen. Nochmals wollen wir festhalten, dass wir die Reserven für den Fall gerechnet haben, wo in einer Gruppe höchstens in 10 % aller Fälle eine grössere Abweichung nach der einen Seite hin auftritt. Die Wahrscheinlichkeit für eine solche Abweichung wäre demnach 0,1, und die Wahrscheinlichkeit dafür, dass diese Abweichungen gleichzeitig bei allen Gruppen auftreten, stellt sich auf $0{,}1^6$. Nun wollen wir aber beachten, dass bei den einzelnen Gruppen die Reservefehler ganz unterschiedlich gross sind. Bei beiden Beständen bleiben je die erste und letzte Gruppe sozusagen ohne Einfluss, während die Abweichung bei der dritten das grösste

Gewicht aufweist. Für eine einzelne Gruppe gesehen, ist es jedoch ohne weiteres möglich, dass auch Abweichungen mit einer kleineren Wahrscheinlichkeit als 10% auftreten. Kommt dies dann gerade in der Gruppe mit dem grössten Gewicht vor, so kann der Fehler bedeutend grösser werden, als wenn in zwei anderen Gruppen zugleich Abweichungen nach der gleichen Seite mit einer Wahrscheinlichkeit von 10% erscheinen. Wir wollen deshalb bei beiden Beständen je in der dritten Gruppe noch eine Wahrscheinlichkeit von 2% zulassen. Das bedeutet dann, dass in dieser Gruppe nur in 1% aller Fälle eine grössere Abweichung als $2{,}326 \times s$ vom $\bar{c}$ der Gruppe nach unten und ebenso nur in 1% aller Fälle nach oben eintritt. Damit glauben wir, dass wir uns gegen alle Überraschungen gesichert haben, denn sollte das gleiche in einer anderen Gruppe vorkommen, dann würde der Fehler ohne weiteres kleiner.

Schlussalter und Reserven für 1% Wahrscheinlichkeiten in Gruppe 3

Bestand	Schlussalter untere 1%-Grenze	Gesamtreserve	Schlussalter obere 1%-Grenze	Gesamtreserve
I	47,70	8 290 288	50,80	8 324 308
Theoretischer I	47,60	8 289 100	50,70	8 322 766
II	50,75	3 331 779	53,40	3 415 432
Theoretischer II	51,00	3 331 732	53,60	3 415 709

Nach dieser Erweiterung erhalten wir folgende Reservefehler:

Bestand I

untere Grenze — 0,217% von der genauen Lidstonereserve
obere Grenze + 0,193% » » » »

Bestand II

untere Grenze — 1,269% » » » »
obere Grenze + 1,210% » » » »

Theoretischer Bestand I

untere Grenze — 0,210% von der Reserve des theoretischen Best.
obere Grenze + 0,196% » » » » » »

Theoretischer Bestand II

untere Grenze — 1,246 % von der Reserve des theoretischen Best.
obere Grenze + 1,244 % » » » » » »

Gesamtbestand I und II

untere Grenze — 0,521 % von der genauen Lidstonereserve
obere Grenze + 0,487 % » » » »

Theoretischer Gesamtbestand I und II

untere Grenze — 0,509 % von der Reserve des theoretischen Best.
obere Grenze + 0,498 % » » » » » »

Die Wahrscheinlichkeit einer grösseren Reserveabweichung käme mit dieser Erweiterung auf $0{,}1^7$ zu stehen; wir wollen aber, um doch etwas vorsichtiger zu sein, je die erste und letzte Gruppe eines Bestandes vernachlässigen, so dass sich die Wahrscheinlichkeit auf $0{,}1^5$ stellt. Wir können dann folgendermassen formulieren: Entnehmen wir jedem Ablaufjahr des Bestandes I eine Stichprobe zu 100 Karten und bilden die Gruppen auf die angegebene Weise, so erwarten wir in höchstens 1 von 100 000 Fällen eine grössere Reserveabweichung als — 0,217 % und höchstens in 1 von 100 000 Fällen eine grössere Reserveabweichung als + 0,193 %, was ganz bestimmt als verblüffendes Resultat zu bezeichnen ist. Das gleiche gilt für Bestand II, nur dass dort die Reserveabweichungen — 1,269 % und + 1,210 % betragen. Die Gründe für diese so viel grösseren Abweichungen haben wir schon früher erläutert. Für den Gesamtbestand erhalten wir durch das Zerlegen in die beiden Bestände höchstens in je 1 von 10^{10} Fällen eine grössere Reserveabweichung als — 0,521 % oder + 0,487 %, was für die Praxis bestimmt als vollständige Sicherheit bezeichnet werden kann. Wir müssen dazu bedenken, dass wir aus dem Gesamtbestand von 23 153 Versicherungen im ganzen 5800 für die Stichproben verwenden mussten, was ziemlich genau einer 25 % Auswahl gleichkommt. Nun wissen wir aber, dass unser Vorgehen insofern nicht vom Umfang der Grundgesamtheit abhängt, als bei dessen Vergrösserung sich die Streuung der Stichproben zu 100 Karten nicht ändern sollte. Damit werden aber auch die Abweichungen der Mittelwerte und die Reservefehler gleich bleiben.

Wie wir früher ausgeführt haben, verwenden wir den theoretischen Bestand als Schätzung der Struktur eines noch grösseren Bestandes, als ihn die Abläufe der 52. Woche des Volksversicherungsbestandes der «Basler Lebens-Versicherungs-Gesellschaft» darstellen. Wir gebrauchen deshalb, wenn wir z. B. die gesamte Reserve dieses Volksversicherungsbestandes berechnen wollten, die Fehler des theoretischen Bestandes und können dann erwarten, dass wir diese Reserve aus 5800 Karten praktisch mit vollständiger Sicherheit mit einem maximalen Fehler von — 0,509 % und + 0,498 % bestimmen können.

§ 5. Versuch aus der Volksversicherung

Als Abschluss dieses Abschnittes haben wir nun noch den Versuch unternommen, aus den Abläufen der 52. Woche des Volksversicherungsbestandes der «Basler Lebens-Versicherungs-Gesellschaft» die Reserven durch Stichproben in der angegebenen Weise zu berechnen, wobei wir erwarten, unter allen Umständen innerhalb der angegebenen Fehlergrenzen zu bleiben.

Stichproben zu 100 Karten mit Gruppenschlussalter

Bestand I					
Gruppe	$\bar{c}$ der Stichproben	Schlussalter	Gruppe	$\bar{c}$ der Stichproben	Schlussalter
1	0,762	48,90	2	0,552	47,40
	0,801			0,947	
	0,918			0,786	
	1,120			0,844	
3	0,881	48,40	4	1,247	52,35
	0,962			1,133	
	0,688			1,301	
	0,832			1,242	
	0,944			1,328	
5	1,349	53,45	6	2,154	56,10
	1,169			1,296	
	1,175			1,142	
	1,753			2,227	
	1,464			2,043	

Die Reserve wird mit diesen Schlussaltern berechnet zu 8 296 263. Die Reserveabweichung beträgt für den Bestand I — 0,145 %, während wir ja als maximale Abweichung — 0,217 % zugelassen haben.

Bestand II					
Gruppe	$\bar{c}$ der Stichproben	Schlussalter	Gruppe	$\bar{c}$ der Stichproben	Schlussalter
1	1,767	55,45	2	1,555	54,45
	1,385			1,261	
	1,808			1,578	
	1,701			1,710	
				1,488	
3	0,969	52,75	4	1,514	53,80
	1,163			1,377	
	1,533			1,433	
	1,354			1,457	
	1,389			1,354	
	1,357				
5	1,757	55,75	6	2,141	57,50
	1,466			2,225	
	1,851			1,925	
	1,729			1,805	
	1,775			2,041	

Die Reserve wird mit diesen Schlussaltern berechnet zu 3 373 944. Für den Bestand II beträgt die Reserveabweichung — 0,019 %, während wir hier als maximale Abweichung — 1,269 % zugelassen haben. Betrachten wir noch die Reserve für den Gesamtbestand I und II, dann erhalten wir als totale Reserve, bestimmt mit Stichproben, den Wert von 11 670 207, während die genaue *Lidstonereserve* 11 682 885 beträgt. Der Fehler für den Gesamtbestand bestimmt sich somit zu — 0,109 %, während wir hier als maximale Abweichung noch — 0,521 % zugelassen haben.

Sämtliche festgestellten Abweichungen bleiben weit innerhalb der zugelassenen Grenzen, so dass uns auch dieser Versuch unsere Überlegungen bestätigt.

V. Abschnitt

Versuch aus der Grosslebensversicherung

Zum Abschluss haben wir noch einen Versuch aus der Grosslebensversicherung unternommen. Wir haben uns dabei als Aufgabe gestellt, für die (nach einem seit 1943 für Abschlüsse geltenden Tarif) in den Jahren 1943–1949 eingegangenen gemischten Versicherungen des Grosslebensbestandes der «Basler Lebens-Versicherungs-Gesellschaft» die Netto-*Lidstonereserve* mit Hilfe eines durch Stichproben bestimmten durchschnittlichen Schlussalters auf die Mitte des Jahres 1950 zu berechnen. Als einzige Einschränkung müssen wir erwähnen, dass wir aus dem Gesamtbestande von 31 879 Versicherungen lediglich 55 Versicherungen mit einer grösseren restlichen Dauer als 40 Jahre und einer gesamten genauen *Lidstonereserve* von Fr. 8114 ausgeschaltet haben (siehe Tabelle auf der nächsten Seite).

Es ist sofort einzusehen, dass wir es hier mit einem ähnlichen Bestand zu tun haben, wie wir ihn schon früher im Bestand II vorfanden. Das ist aber für unsere Aufgabe gerade der allerungünstigste Fall, denn es fehlen die grossen Reserven kurz vor Ablauf fast vollständig. Aus den Schlussaltern ist ersichtlich, dass wir auch hier die Gruppeneinteilung vornehmen können, denn die Schlussalter einer Gruppe stimmen eher noch besser als in den früher besprochenen Beständen überein. Wir haben aus der zweiten bis sechsten Gruppe für jedes Ablaufjahr je eine Stichprobe zu 100 Karten entnommen, aus der ersten und siebenten aber nur je Stichproben zu 50 Karten, da dort die Zahl der vorkommenden Versicherungen noch sehr klein ist. Mit diesen Stichproben haben wir wieder das Gruppenalter bestimmt und damit die Reserve gerechnet. Wir müssen uns bei diesem Versuch klar darüber sein, dass wir keine näheren Anhaltspunkte über die Streuung der Stichproben zu 100 Karten besitzen, so dass wir über die Wahrscheinlichkeiten von Reserveabweichungen kaum etwas aussagen können. Da aber in diesem Bestande die Kinderversicherungen, ohne welche die Streuung bestimmt kleiner ausfallen wird, fast vollständig fehlen, dafür aber viel grössere Unterschiede in den Versicherungssummen auftreten, wodurch die Streuung wieder grösser ausfallen wird, so glauben wir doch, dass hier die Streuungsverhältnisse nicht

Grosslebensbestand mit Schlussalter und genauer Lidstonereserve

Ablaufjahr	Anzahl Versicherungen	Schlussalter	Genaue Lidstonereserve	Ablaufjahr	Anzahl Versicherungen	Schlussalter	Genaue Lidstonereserve
1953	75	58,75	350 436	1970	1696	58,35	1 818 569
1954	132	60,85	525 715	1971	1572	58,70	1 069 618
1955	170	61,25	548 090	1972	1511	59,20	879 951
1956	188	61,15	526 156	1973	1956	58,25	1 022 303
1957	210	61,60	658 379	1974	1686	58,05	782 507
1958	490	61,05	1 140 302	1975	1139	59,25	729 566
1959	531	61,65	1 085 494	1976	1175	58,85	623 167
1960	511	60,60	875 674	1977	1138	58,05	455 050
1961	610	62,00	1 019 246	1978	1058	58,10	282 414
1962	632	62,10	927 988	1979	808	58,95	157 411
1963	1468	59,10	2 095 407	1980	348	60,25	130 197
1964	1615	60,15	2 084 545	1981	326	61,50	115 191
1965	1647	60,25	1 879 840	1982	285	61,35	89 075
1966	1669	59,20	1 520 624	1983	261	62,45	49 002
1967	1591	58,95	1 137 976	1984	183	62,75	38 855
1968	2499	58,70	2 012 018				
1969	2194	58,70	1 378 153	1985	133	62,45	29 898
				1986	120	62,85	26 957
				1987	87	62,85	15 553
				1988	67	63,50	5 631
				1989	43	62,90	— 531
				Gesamte *Lidstonereserve:*			27 586 427

allzusehr von denen der früheren Bestände abweichen werden. Wir können somit etwa erwarten, dass der Reservefehler innerhalb der Grenzen von $\pm 1{,}2\,\%$ bleiben wird. Weiter sei erwähnt, dass wir bei der Stichprobenerhebung keinerlei Vorsichtsmassnahmen getroffen haben, um beispielsweise Versicherungssummen von Fr. 50 000 oder Fr. 100 000, die das Schlussalter empfindlich stören können, auszuschalten. Wir haben dann auch wirklich in mehreren Stichproben solch grosse Summen gezogen.

Stichproben zu 100 Karten mit Gruppenschlussalter

Jahr	Anzahl Karten	$\bar{c}$ der Stichproben	Schluss-alter	Jahr	Anzahl Karten	$\bar{c}$ der Stichproben	Schluss-alter
1953	50	2,063	60,40	1970	100	2,228	59,05
1954	50	2,413		1971	100	2,643	
1955	50	2,927		1972	100	2,075	
1956	50	2,385		1973	100	2,560	
1957	50	3,501		1974	100	2,225	
1958	100	2,787	61,20	1975	100	2,251	59,35
1959	100	2,754		1976	100	2,763	
1960	100	2,719		1977	100	2,387	
1961	100	2,769		1978	100	2,277	
1962	100	3,296		1979	100	2,339	
1963	100	2,104	58,80	1980	100	2,711	61,75
1964	100	2,370		1981	100	2,892	
1965	100	2,637		1982	100	2,943	
1966	100	2,208		1983	100	3,241	
1967	100	2,010		1984	100	3,307	
1968	100	2,422					
1969	100	2,281		1985	50	2,938	62,40
				1986	50	3,323	
				1987	50	3,357	
				1988	50	3,076	
				1989	43	3,360	

Wir erhalten als Stichprobenreserve 27 648 943 und als genaue *Lidstonereserve* 27 586 427, also eine Abweichung von $-0{,}227\,\%$, die somit weit innerhalb der erwarteten Grenze liegt.

Auch dieser Versuch bestätigt uns unsere Überlegungen, und das Verfahren kann bestimmt bereits als rationell angesprochen werden, haben wir doch aus dem Gesamtbestand von 31 824 Versicherungen nur 3193 Versicherungen für die Stichproben verwendet, was ziemlich genau einer 10% Auswahl entspricht.

Zusammenfassung

In der vorliegenden Arbeit untersuchten wir, ob nicht bei Wahl der Methode von *Lidstone* zur Ermittlung des Deckungskapitals an Stelle des genauen ein durch Stichproben bestimmtes durchschnittliches Schlussalter verwendet werden könnte. Wir erläuterten die Berechtigung dieser Fragestellung im ersten Abschnitt. Im zweiten Abschnitt erwähnten wir die für die Arbeit benötigten Prüfverfahren, und im dritten Abschnitt belegten wir die Normalverteilung der Stichproben durch Versuche. Im vierten Abschnitt arbeiteten wir ein Verfahren aus, mit dem ein durchschnittliches Schlussalter durch Stichproben günstig bestimmt werden kann. Im gleichen Abschnitt legten wir für unsere Untersuchungsbestände Reservefehler fest, die kaum mehr überschritten werden dürften, und bewiesen unsere Überlegungen durch einen Versuch aus der Volksversicherung. Im fünften Abschnitt endlich nahmen wir einen Versuch bei Grosslebensversicherungsbeständen vor, um auch dort die Berechtigung unserer Überlegungen zu zeigen.

Allgemeingültig können wir auf Grund unserer Arbeit aussagen, dass bei normalen Versicherungsbeständen mit Hilfe weniger Stichproben eine überraschend gute Übereinstimmung erzielt werden kann, dass aber auch bei Neubeständen der maximale Fehler kaum grösser als derjenige ist, mit dem die Methode der Reserverechnung nach *Lidstone* in bezug auf das genaue Einzelverfahren ohnehin schon behaftet ist. Wenn wir aber für einen anderen als den untersuchten Versicherungsbestand zuverlässige Grenzen der Abweichungen bestimmen wollen, werden wir kaum darum herum kommen, die Grösse der Streuung der Stichproben wenigstens einmal zu bestimmen.

Literaturverzeichnis

[1] *E. Zwinggi:* Versicherungsmathematik, Basel, 1945.

[2a] *G. C. Campbell:* Problems with Sampling Procedures for Reserve Valuations. Journal of the American Statistical Association, Vol. 43, 1948.

[2b] *E. Zwinggi:* The Calculation of Reserves by Means of Samples. Schweizerische Zeitschrift für Volkswirtschaft und Statistik. 85. Jahrg., 1949.

[2c] *E. Zwinggi:* Anwendung neuerer statistischer Verfahren in der Versicherungsmathematik. Blätter der Deutschen Gesellschaft für Versicherungsmathematik, 1950.

[3a] *M. G. Kendall:* The Advanced Theorie of Statistics, Vol. I, II, London, 1943, 1946.

[3b] *R. A. Fisher:* Statistical Methods for Research Workers, London, 1946.

[3c] *A. Linder:* Statistische Methoden, Basel, 1945.

[4] *R. v. Mises:* Wahrscheinlichkeitsrechnung, Wien, 1931.

[5] *H. L. Rietz-Baur:* Handbuch der mathematischen Statistik, Berlin, 1930.

[6] *K. Pearson:* On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling. Phil. Mag. 5, Vol. 50, 1900.

[7] *R. A. Fisher:* Applications of «Student's» distribution. Metron Vol. 5, No. 3, 1926.

[8] *R. A. Fisher:* The moments of the distribution for normal samples of measures of departure from normality. Proceedings of the Royal Society London. Vol. 130 A, 1931.

[9] Siehe [3a] Vol. I, p. 280.

[10] *E. S. Pearson:* A further development of tests for normality. Biometrica Vol. 22, 1930.

[11a] Siehe [3a] Vol. II, p. 106.

[11b] *R. A. Fisher:* The conditions under which χ^2 measures the discrepancy between observation and hypothesis. Journal of the Royal Statistical Society. Vol. 87, 1924.

[12] Siehe [3a] Vol. II, p. 175.

[13] *W. F. Sheppard:* The calculation of the moments of a frequency distribution. Biometrica Vol. 5, 1907.

Lebenslauf

Ich wurde am 13. November 1925 als Sohn des Ernst Baumann, Bautechniker, von Basel, und der Olga, geb. Neff (gest. 1947), in Paris geboren. Von 1932–1936 besuchte ich in meiner Heimatstadt Basel die Primarschule und trat im Frühjahr 1936 in das Humanistische Gymnasium in Basel ein. Im Herbst 1940 wechselte ich ins Mathematisch-naturwissenschaftliche Gymnasium Basel, wo ich im Frühjahr 1945 die Maturität bestand. Anschliessend immatrikulierte ich mich an der Philosophisch-naturwissenschaftlichen Fakultät der Universität Basel, um mich vor allem dem Studium der Versicherungsmathematik, Mathematik, Nationalökonomie und Experimentalphysik zu widmen.

Den Herren Professoren Baldinger, Buchner, Furlan, Gürtler, Huber, Miescher, Ostrowski, Renfer, Salin, Speiser, Wagner und Zwinggi danke ich für die Anregungen, die ich beim Besuch ihrer Vorlesungen, Übungen und Seminarien empfangen habe.

Die vorliegende Arbeit ist unter der Anleitung von Herrn Prof. Dr. Zwinggi entstanden; ihm bin ich für die mir während meines ganzen Studiums gewährte Unterstützung zu besonderem Dank verpflichtet.

Weiter gebührt mein Dank der Direktion der «Basler Lebens-Versicherungs-Gesellschaft» für die Überlassung des Materials und die bei der Verarbeitung gewährte Hilfe.

Robert Baumann